# SILICON GRAPHICS, INC.
# ORIGIN 2000 CLUSTER
# USER'S GUIDE

# ERDC DoD HIGH PERFORMANCE COMPUTING
# MAJOR SHARED RESOURCE CENTER

**VERSION 1.0**
**September 1998**

**US Army Engineer Research and Development Center, Vicksburg, MS**

# Preface

The US Army Engineer Research and Development Center (ERDC) in Vicksburg, Mississippi is the premier research and development laboratory complex of the Corps of Engineers. The Army Supercomputer Center at ERDC was established in 1989. In 1993, ERDC began operations as the first of the Department of Defense (DoD) High Performance Computing (HPC) Major Shared Resource Centers (MSRC). The ERDC MSRC was formed under the auspices of the DoD HPC Modernization Program and is within the Information Technology Laboratory (ITL) at ERDC. The ERDC MSRC mission is to provide state-of-the-art scientific HPC support to all Defense Research, Development, Test, and Evaluation (RDT&E) user communities. Access to the ERDC MSRC systems is available through multiple, nationwide high-speed data communications networks. In addition, training and consultation is provided to local and remote DoD users of these systems.

Questions, comments, and suggestions about this guide are welcomed. Comments may be sent in the following ways:

| | |
|---|---|
| Toll-free long distance: | 800-500-HPCC (4722) |
| Local WES telephone: | 601-634-4400, option 1 |
| E-Mail: | info-hpc@wes.hpc.mil |
| Facsimile: | ERDC Customer Support (ATTN: Margaret Valentine)<br>601-634-3808 |
| U.S. Postal Mail: | US Army Engineer Research and Development Center<br>ATTN: ERDC-IM-H (Margaret Valentine)<br>3909 Halls Ferry Road<br>Vicksburg, MS 39180-6199 |

The ERDC MSRC also maintains a site on the World Wide Web (WWW) at the Unified Resource Locator (URL) http://wes.hpc.mil. Users requiring printed copies of Center documentation may contact the ERDC MSRC Customer Support.

# Table of Contents

Preface ........ i

Table of Contents ........ ii

Release History for ERDC MSRC Silicon Graphics Origin 2000 Cluster User's Guide ........ iv

1. Introduction ........ 1-1
    - Scope 1-1
    - Font Conventions ........ 1-1
    - System Configuration ........ 1-1
    - System Access ........ 1-2
    - ERDC Customer Support ........ 1-2
    - World Wide Web Information ........ 1-3
    - Obtaining an Account ........ 1-3
    - Security ........ 1-4
    - Services and information ........ 1-4
    - Training ........ 1-5

2. System Resources ........ 2-1
    - Hardware ........ 2-1
    - System Utilization/Efficiency ........ 2-2
    - Data Storage ........ 2-2
        - Permanent File Storage ........ 2-2
        - Temporary File Storage ........ 2-4
            - Managing Temporary File Storage ........ 2-4
    - Network Connectivity ........ 2-5
    - Application Support Software ........ 2-5
        - Applications/Utilities File Systems ........ 2-5

3. Program Development ........ 3-1
    - Overview of Compilers and Development Tools ........ 3-1
        - Summary of Available Compilers and Compiler tools ........ 3-1
    - C and C++ Compilers ........ 3-2
    - FORTRAN 77 and FORTRAN 90 ........ 3-3
    - Parallelizing Compilers ........ 3-3
        - Power C Analyzer ........ 3-3
        - Power FORTRAN Accelerator ........ 3-4
        - High Performance Fortran ........ 3-4
            - pghpf File Extension Summary ........ 3-5

4. Message Passing Models....................................................4-1
Overview....................................................4-1
Message Passing Interface Model....................................................4-1
Parallel Virtual Machine Model....................................................4-2

5. Programming Environment....................................................5-1
Batch Request Environment....................................................5-1
Interactive Environment....................................................5-3

# Release History for ERDC MSRC O2K User's Guide

**Release 1.0**

September 1998
Original Release

# 1. Introduction

## Scope

This document provides an overview and introduction to the use of the Silicon Graphics, Inc. (SGI) Origin 2000 (O2K), located at the ERDC MSRC, and a description of the specific computing environment on the O2K. The intent of this guide is to provide information that will enable the average user to perform computational tasks on the O2K system. To receive the most benefit from the information provided here, the user should be proficient in the following areas:

- Use of the UNIX operating system..
- Use of an editor (e.g., vi or emacs).
- Remote usage of computer systems via network or modem access.
- A selected programming language and its related tools and libraries.
- General information about the ERDC MSRC (found in the *ERDC MSRC Introductory User's Guide*).

## Font Conventions

The following font conventions will be used in this manual:

| | | |
|---|---|---|
| **boldface** | commands, system prompts | (e.g., **ls -l**) |
| ***boldface italic*** | user response | (e.g., **login:*my_user_name***) |
| *italic* | important information | (e.g., *tape backups*) |
| `courier font` | filenames, hostnames, output, and shell environment variables | (e.g., `/etc/motd`) |
| <cr> | carriage returns, "enter" key | |

## System Configuration

The ERDC MSRC O2K is comprised of an SGI Origin 2000 shared memory multiprocessor system, called origin. The system contains 128 central processing units (CPUs) divided into a cluster of two operating system images. The first image contains sixteen 195 MHz MIPS R10000 CPUs and 8 gigabytes (GB) of RAM. The second image (called sdf1) contains 112 CPUs and 56 GB of RAM. The Origin cluster contains over 600 GB of formatted disk space, with 417 GB set aside as a temporary user work area. For instructions in the use of temporary storage please see page 2-4 of this document.

# System Access

To access the ERDC MSRC O2K, **telnet** to `origin.wes.hpc.mil`.

Toll-free modem access to the ERDC MSRC O2K is also available. The number is 800-500-9472 with suggested settings of 8 data bits, 1 stop bit, no parity, and up to 28.8 Kb/sec. Use of the modem requires a dial-in account, which may be obtained from Customer Support. Upon connection to the modem, use the **telnet** command as described above.

Users are required to periodically change their passwords for obvious security reasons. To accomplish this goal, passwords are "aged" on the ERDC MSRC systems. The user can initiate a password change for their own account by typing **passwd**. Please see the additional password-related information under "Obtaining an Account," page 1-3 and the Kerberos section, page 1-4.

# ERDC Customer Support

The ERDC MSRC provides application support for the following Computational Technology Areas (CTA):

- Computational Structural Mechanics (CSM)
- Computational Fluid Dynamics (CFD)
- Environmental Quality Modeling and Simulation (EQM)
- Climate/Weather/Ocean Modeling (CWO)
- Forces Modeling and Simulation (FMS)

ERDC Customer Support provides assistance with problems, questions, training requests, as well as other user-related services. Customer Support analysts are available Monday through Friday from 6:00 AM until 7:00 PM, Central Time (excluding government holidays). Outside these times, the toll-free user assistance lines are connected to an answering service and calls will be returned the next business day. After hours user assistance telephone calls to 601-634-4400, option 1, are automatically routed to the ERDC ITL computer operations staff, which is available 24 hours per day, including weekends and holidays. For local users, ERDC Customer Support is located in TL-112, room 105. Please see Preface, page i, of this document for telephone numbers.

# World Wide Web Information

Information concerning the facilities and system updates, as well as user's guides and account request forms, are available at the ERDC MSRC web site. Users are encouraged to visit the site often as new information is added. Please see Preface, page i, of this document for site information.

In an effort to better serve the user community, the computing environment on the O2K is subject to change. Updated documentation will be made available as changes are implemented.

# Obtaining an Account

Authorized DoD and contractor personnel may obtain user accounts on the ERDC MSRC systems by filling out a "DoD HPC Project and Account Application Form," available at the ERDC MSRC web site. Forms may also be obtained from the designated DoD HPC Service/Agency Approval Authority (S/AAA) or Terminal Area Security Officer (TASO). DoD and contractor personnel obtain authorization by completing Section Two of the form (user information) and forwarding the form to their S/AAA. The local S/AAA is responsible for providing the project information required in Section One of the form, and returning the completed form to ERDC. Please contact the HPCMO home page (http://www.hpcmo.hpc.mil/Htdocs/SAAA/saaa.html) or call Customer Support to obtain S/AAA and TASO information if necessary.

Each new user will also be required to sign an account receipt form which includes a security statement. **SHARING OF USERIDS/PASSWORDS IS SPECIFICALLY PROHIBITED ON ALL ERDC MSRC HPC SYSTEMS.**

Policy requires that everyone accessing the ERDC MSRC computer systems must have a National Agency Check (NAC). The general procedures are as follows:

Program Environment and Training (PET) members and Nichols Research Corp. (NRC) employees. Contact Ms. Jeannie McDonald at 1-601-634-2609.

WES employees and WES contractors. Contact Ms. Linda McGowan, WES Security Office at 1-601-634-2776.

Outside WES. Individuals should contact their local security office for a NAC packet. Fingerprints can be taken at any local law enforcement agency, but must be on the official fingerprint card included in the NAC packet.

Any further questions should be directed to Mr. William H. Roth, Jr., Security Officer, WES, 1-601-634-3663, or Ms. McGowan at 1-601-634-3527.

# Security

The ERDC MSRC is implementing a three-phase transition to a secure computing environment at this time. The overlapping transition began with the implementation of Secure Shell, and has been enhanced by Kerberos, Version 5. A transition to Distributed Computing Environment and the addition of Distributed File System will complete the secure environment.

## Using Kerberos

### Procedure To Login To ERDC MSRC HPC Systems With Kerberos/SSH

To obtain a KRB5 ticket on a UNIX-based system:

```
%kinit
Password for user@WES.HPC.MIL: enter Kerberos password
Challenge for Security Dynamics mechanism: []
Passcode: enter your PIN number into the Kerberos card, press the diamond,
and then enter the six-digit passcode here.
```

To verify that you have a received a Kerberos ticket:

```
%klist
Ticket cache: /tmp/krb5cc_982
Default principal: user@WES.HPC.MIL

Valid starting          Expires          Service principal
07 Jun 98 10:32:2008 Jun 10:32:20krbtgt/WES.HPC.MIL@WES.HPC.MIL
```

3. To change your KRB5 password:

```
%/usr/local/bin/kpasswd
kpasswd: Changing password for user@WES.HPC.MIL
Old password:
Challenge for Security Dynamics mechanism: []
Passcode:
New password:
New password (again):
Kerberos password changed.
```

On a workstation which has been Kerberized (has the client sofware installed in `/usr/local/bin`) **use /usr/local/bin/telnet** as the Kerberized telnet.

a. To access a KRB5 system, use **telnet systemname**<br>
b. To a non-Kerberized system, use **ssh systemname**

### Using Kerberos Telnet

If on a workstation, do the following:

1. The user must first establish a Kerberos ticket by typing `/usr/local/bin/kinit`. The user will be prompted to enter their Kerberos password. After entering the Kerberos password, the user will be prompted to enter a passcode. A passcode is the six digit number displayed on the user's SecurID card after the PIN is entered and the diamond key is pressed. The SecurID card always displays a six digit number that changes periodically. To obtain a valid passcode the user must type the PIN for his/her SecurID card on the keypad of the SecurID card and press the diamond. This will cause the SecureID card to issue a valid passcode. The user should type this passcode into the terminal. The system will then return a prompt.

2. Once a Kerberos ticket has been granted, the user can establish a **telnet** session to the desired remote host, provided that the host has Kerberos Version 5 installed and running, by typing **/usr/loca/bin/telnet <HOSTNAME>**.

3. For instructions on using Kerberos with a PC, contact Customer Assistance.

## Services and Information

Users of the ERDC MSRC systems are provided with information through the toll-free Customer Support hotline, workshops and seminars, the WWW site, and on-line documentation. A brief discussion of some on-line services follows:

- Message of the Day - An informative message of the day (**motd**) is displayed upon login to the ERDC systems. The **motd** contains important information about imminent events that will affect the immediate usage of the system. The UNIX **pg** command is used with **motd** to prevent longer messages from scrolling off the monitor. The message is located in the file `/etc/motd` and may be viewed at any time by issuing the command **pg /etc/motd** at the system prompt. Please read this information carefully.

- On-Line Bulletin System. An on-line bulletin system is available on each system and can be used to obtain important information about the system. The bulletins are usually brief and contain information on a variety of topics. To display the list of available bulletins, type:

  **bull**

  A menu with a list of available bulletins will be displayed on the screen. Press <return>, enter the number of the bulletin you wish to display, and press <return> again. The bulletin will be displayed at your terminal one screen at a time. Press

<spacebar> to display each additional screen until you reach the end of the bulletin. Press **q** and <return> to exit the bulletin utility from the main menu. For more information about the **bull** command, type:

```
man bull
```

# Training

The ERDC MSRC also supports an extensive training schedule through its Programming Environment Training (PET) function. Most training is conducted at the ERDC MSRC Training and Education Facility (TEF) in the Information Technology Laboratory (ITL) of ERDC. Training at remote facilities and/or specialized training courses will be considered upon request. Please contact Customer Support for more information or to submit training requests. Also, for users migrating from the Cray C90 to the O2K the Computational Migration Group (CMG) is available for help. The training schedule is updated regularly on the ERDC MSRC WWW page. Please see Preface, page i, of this document for URL and telephone numbers.

# 2. System Resources

## Hardware

The O2K consists of 128 MIPS R10000 195 MHz processors and 64 GB of memory in a cache-coherent non-uniform memory access (ccNUMA) distributed shared memory (DSM) architecture. Each processor contains a 32 GB primary data cache, a 32 KB primary instruction cache, and a 4 MB secondary cache.

The O2K is formed into a cluster of two separate operating system images; the first consisting of 16CPUs with 8 GB RAM, and the second consisting of 112 processors with 56 GB RAM. The cluster has an aggregate peak computational speed of 50 GFLOPS.

The Origin cluster is accessed by logging into the 16-CPU image, called `origin`. User home directories are served from another O2K, which is configured as a high availability file server (HAFS). HAFS has 214 GB of home directory space available, with a current quota of 100 MB per user.

Production jobs on the Origin cluster are run in batch mode on the 112-CPU machine, named `sdf1`, which sports 417 GB of available local disk space for user workspace. Users submit jobs from host `origin`, and those jobs then execute on host `sdf1` as resources become available.

The work space area is named `/Work` on the Origin cluster. This filesystem is local to `sdf1` and is NSF mounted to `origin`. Users should access this filesystem in their scripts by using the shell environment variable $WORKDIR, which expands to `/Work/<username>`. Users who are accustomed to using `/tmp` for workspace on other systems are advised not to use `/tmp` on the Origin cluster. `/Work` is the correct temporary workspace area. Files remaining on `/Work` for longer than 15 days are subject to removal when space becomes limited. Always make backup copies of working files and place them on the mass storage system to ensure safe keeping.

## System Utilization/Efficiency

Users should not interactively execute their program or job on `origin.wes.hpc.mil`. The host `origin` is for compiling, editing, and general interactive use by all CEWES MSRC

Origin cluster users. In the interest of the rest of the users logged in, jobs found running on host `origin` may be unilaterally terminated. The preferred method to run interactively is to use PBS. The correct command is **qsub -I**. Specify the number of cpus and the required wall time. Jobs submitted in this manner will be submitted to `sdf1` for execution.

# Data Storage

## Permanent File Storage

On the O2K, each user is allocated a home directory (the current working directory immediately after login) with an initial disk quota of 100 MB of permanent non-migrated storage. The user's home directory can be referenced locally with the $HOMEDIR environment variable.

Requests to increase disk space quotas or interactive limits may be submitted by contacting Customer Support for a special request form. The following information must be supplied by the user for evaluation of the request by the system administrators and CEWES MSRC management:

- the amount of system resource requested
- the length of time requested for the increase
- special deadlines for the project
- an explanation of the attempts to work within limits

All MSRC systems share an on-line Mass Storage Facility (MSF) which currently includes over 200 GB of high-speed disk storage and over 75 TB of high-speed robotic archival storage utilizing a StorageTek 9310 robotic silo. The MSF should be used for all long-term storage (longer than 90 days).

Every O2K user has an account on the MSF. A user's MSF home directory can be referenced from the O2K or the MSF by using the $MSFHOME environment variable from any shell environment. In addition, the **msfstat**, **msfput**, and **msfget** commands may be used for checking the on-line status of the MSF and transferring files to and from the MSF. The following example batch mode script utilizes these commands to transfer files to and from permanent storage on the MSF:

```
#!/bin/ksh
# This is a sample PBS script
#
# Export login shell environment to PBS
#PBS -V
#
# Merge stdout and stderr streams into a single output file
#PBS -j oe
```

```
#
# Ask for 48 CPUs
#PBS -l ncpus=48
#
# Request 1 hour of wallclock time for execution
#PBS -l walltime=01:00:00
#
# Change directory to home directory and copy
# input file to working storage
```

```
cd $HOMEDIR
cp input.dat $WORKDIR
#
# Change directory to working storage space
cd $WORKDIR
#
# Check that mass storage facility is available;
# if not, wait in quiet mode (no output messages)
msfstat -q -w
#
# Copy file "test.exe" from MSF in user's subdirectory
"source"
msfget $MSFHOME/source/test.exe $WORKDIR
#
# execute program
test.exe > test.out
#
# recheck availability of MSF before attempting to move data
msfstat -q -w
#
# copy output file to MSF
msfput test.out $MSFHOME/output/test.out
#
# clean up files no longer needed
/bin/rm test.exe test.out
```

Although the MSF provides long-term storage for user data, users are cautioned that system backups are made for disaster recovery only and backup tapes are periodically recycled. Consequently, they are not available for user file restoration and users are **strongly encouraged** to make tape backups of critical data. For complete information regarding the use of permanent file storage, refer to the *CEWES MSRC Mass Storage Facility Introductory User’s Guide*.

## Temporary File Storage

417 GB of temporary storage is available to all O2K users and is intended for the *temporary* storage of data files needed for executing programs. Users should access their temporary storage space via the use of the $WORKDIR environment variable. This variable is set for every user upon login. Files located in $WORKDIR do not affect the user’s permanent file quota usage. Interactive and batch jobs are allowed to use large amounts of disk storage as there are no quota limits on the amount of disk space used in $WORKDIR. This fact, multiplied by the large number of CEWES MSRC users, predisposes the filesystem to space shortages.

## Managing Temporary File Storage

Close management of the user’s temporary storage is a very high priority, because when disk space becomes too low, the system halts processing and manual intervention is required to restart processing. Users are responsible for managing their own files in $WORKDIR through the transfer of needed files to MSF and deletion of unneeded files when their processes end. Orphan files (files that are not owned by an active process) are allowed to remain in $WORKDIR for fifteen days unless the filesystem becomes low on available space. Users who leave orphan files in $WORKDIR, or use commands such as **touch** to simply freshen their files, are considered misusers.

# Network Connectivity

The internal MSRC networks are built on redundant combinations of Fiber Optic Distributed Data Interface (FDDI), High Performance Parallel Interface (HiPPI), high speed ethernet, and Asynchronous Transfer Mode (ATM) network technologies. Additionally, CEWES MSRC has direct, redundant connectivity to the Defense Research and Engineering Network (DREN) network. The O2K can be reached via network connection to `origin.wes.hpc.mil`. Use of the hostname is preferred since Internet Protocol (IP) addresses are subject to change

# Application Support Software

All CEWES MSRC systems run derivatives of the UNIX System V operating system with vendor-specific enhancements. A large variety of compiler environments, math libraries, graphics libraries, and third-party analysis applications are available on the systems. Many more applications are being added to accommodate the diverse needs of the user communities served by the CEWES MSRC. The following application software packages are or will be available on the O2K:

| Product | Version |
|---|---|
| TOTALVIEW Multiprocess Debugger | Latest Version |
| VAMPIR MPI Profiling Library | Latest Version |
| ABAQUS Std/Exp/Post/AQUA/USA | 5.7.x |
| ENSIGHT | 6.1 |
| pgHPF Portland Group High Performance Fortran | 2.4-4 |
| CSA/NASTRAN (Network) | Latest Version |
| NCAR Graphics | 4.0.1 |
| FAST | Latest Version |
| VAST-90 Fortran 90 Compiler/Fortran 77/90 Translator | 3.4k/4.3c |

## Applications/Utilities File Systems

In addition to the software mentioned previously, applications and utilities on the O2K include SGI-proprietary and supported programs, commercial packages, and privately-written and supported programs. All non-SGI applications software, libraries, utilities, and documentation are stored on each CEWES MSRC system in the `/usr/local` directory structure. Users should consult the following important subdirectories of `/usr/local` to find the applications and utilities listed below:

`/usr/local` All non-SGI applications software, libraries, utilities, and documentation.

`/usr/local/applic` Third-party applications.

`/usr/local/bin` Third-party software executable files and system shell scripts.

`/usr/local/doc` Limited documentation of third-party software, libraries, and utilities.

`/usr/local/info` Bulletin files for user access.

`/usr/local/man` Locally-developed CEWES MSRC **man** pages.

`/usr/local/usp` Contents of these subdirectories are not supported by CEWES MSRC systems personnel, but rather by the owner/user. Subdirectories named bin, lib, applic, and org contain the programs.

`/usr/local/usp/org` User site-specific directories can be established within this directory for software needed only by a particular organization rather than by the entire user community. User software must meet certain requirements to qualify for placement in this directory. Contact Customer Support for information on this topic.

# 3. Program Development

## Overview of Compilers and Development Tools

The CEWES MSRC O2K provides a full complement of programming development tools. These tools include assemblers, compilers, parallelizing compilers, and programming utilities. The following sections describe these elements of the O2K programming environment.

More information about these compilers can be found by using the online "man pages" or online SGI Insight Books. Man pages can be seen by typing **man [command]**. Insight Books, an X-based program, can be accessed by simply typing **insight** after setting the environment variable $DISPLAY to point to your local host. They are also available on the web.

Summary of Available Compilers and Compiler tools

| COMMAND | COMPILER |
|---|---|
| **as** | Assembler |
| **abicc** | ABI C Compiler |
| **CC** | C++ Compiler |
| **cc** | C Compiler |
| **f77** | FORTRAN 77 Compiler |
| **f90** | FORTRAN 90 Compiler |
| **NCC** | 32-bit Delta/C++ Compiler, see page 3-2 for details |
| **DCC** | 32-bit Delta/C++ Compiler, see page 3-2 for details |
| **pghpf** | Portland Group High Performance FORTRAN (PGHPF) |
| | **PREPROCESSORS** |
| **cpp** | C Language Preprocessor |
| | **TRANSLATION TOOLS** |
| **mpc** | Multiprocessing C Source Transformer |

# C and C++ Compilers

The O2K provides compilers for both the C and C++ programming languages. The C compiler conforms to the ANSI C standard as well as "traditional C," the dialect of C defined by Kernigan and Ritchie in *The C Programming Language*. Compiler options allow compilation of programs written in "traditional C," pure ANSI C, or ANSI C with SGI extensions.

The O2K implementation of C++ conforms to the standard as defined in *The Annotated C++ Reference Manual*, by Margaret Ellis and Bjarne Stroustrup. The IRIX operating system supports the following two versions of the C++ compiler:

- NCC - This is a 32-bit native compiler that implements all the features of the language described in *The Annotated C++ Reference Manual*.

- DCC - This is the Delta/C++ compiler that is available as part of the CASEVision/Workshop Pro C++ Package. Delta C++ is an extension to C++. It is a native compiler that supports dynamic classes, which minimizes the need for recompilation if classes are modified.

The C and C++ command-line syntax is given below:

**cc [option(s)] filename[...]**
or **CC [option(s)] filename[...]**

Where

**option**(s) are one or more command-line options.

**filename** is the name of the source file, assembly-language file, object file, or library to be processed by the compilation system. More than one filename may be specified.

| **Filename Extension** | **Assumed Type** |
|---|---|
| `.c` | C source programs |
| `.C` | C++ source programs |
| `.s` | Symbolic assembly language source |
| `.i` | File containing output from the source preprocessor |

After compilation, an extension of '`.o`' will be added to each object program produced. For further information and a list of options, use **man cc**, **man CC**, or **insight**.

# FORTRAN 77 and FORTRAN 90

The full ANSI Programming Languages FORTRAN 77 and FORTRAN 90 are available on the O2K with a comprehensive set of FORTRAN extensions.

The FORTRAN 77 and FORTRAN 90 command-line syntax is given below:

**f77 [option(s)] filename[...]**
or **f90 [option(s)] filename[...]**

Where

**option**(s) is one or more command-line options, see below.

**filename** is the name of the source file, assembly-language file, object file, or library to be processed by the compilation system. More than one filename may be specified.

| **Filename Extension** | **Assumed Type** |
|---|---|
| `.f`, `.F`, or `.f90` | FORTRAN 77 or FORTRAN 90 source programs |
| `.s` | Symbolic assembly language source |
| `.i` | File containing output from the source preprocessor |

After compilation, an extension of '`.o`' will be added to each object program produced. For further information and a list of options, use **man f77, man f90,** or **insight**.

# Parallelizing Compilers

## Power C Analyzer

The IRIX Power C Analyzer is a parallelizing preprocessor that detects potential parallelism in C code. It also performs other optimizing tasks such as the following:

- Determines data dependencies, which might prevent code from running in parallel.
- Distributes well-behaved loops and certain other code across multi-processors.
- Optimizes source code.

The Power C Analyzer is invoked as an option to **cc**/**CC**. The command-line syntax is given below:

**cc -pca [options] filename [...]**
or **CC -pca [options] filename [...]**

For further information, use ***man cc***, or ***insight***.

## Power FORTRAN Accelerator (PFA)

The Power FORTRAN Accelerator is a parallelizing preprocessor that attempts to optimize FORTRAN programs to run in parallel on the O2K. PFA analyzes a program and identifies loops without data dependencies. It is a preprocessor that automatically inserts special compiler directives into a FORTRAN program to produce a modified copy of the source. The FORTRAN compiler can then interpret these directives to generate code that can run across multiple processors. PFA does not affect the portability of the code to non-SGI systems.

The Power FORTRAN Analyzer is normally invoked as an option to f77/f90. The command-line syntax is given below:

```
     f77 -pfa [options] filename [...]
or   f90 -pfa [options] filename [...]
```

For further syntax definition and an explanation of the available options, use **man f77, man f90**, or **insight** to reference the *Power FORTRAN Accelerator User's Guide*.

## High Performance Fortran (HPF)

The Portland Group High Performance FORTRAN provides data parallelism Fortran. It also adds the HPF Library, which provides optimized intrinsics for sorting, scanning and reduction operations on arrays. An "extrinsic" capability allows a programmer to make calls to existing subroutines that were written using message-passing or other parallel models.

**pghpf** bases its processing on the suffixes of the files it is passed. Files whose names end with `.hpf`, `.f`, `.for`, or `.f90` are taken as HPF source files. They are compiled and assembled. Files whose names end with `.F` are also considered HPF source files, but are first processed by the C preprocessor. Files whose names end with `.s` are considered to be assembly language files. Files whose names end with `.o` are taken as object files, and are passed directly to the linker if linking is requested. Files whose names end with `.a` are taken as libraries. No action is performed on `.a` files unless linking is requested. Files not ending in `.o`, `.s`, `.f`, `.f90`, `.for`, `.F`, `.hpf`, or `.a` are taken as object files and passed to the linker (if linking is requested) with a warning message. Unless the default **pghpf** action is overridden using a command-line option, **pghpf** deletes intermediate FORTRAN 77, assembly, and object files (see option **Mkeepftn**). The **pghpf** command-line syntax is given below:

```
pghpf [options], filename [...]
```

To run a pghpf program: a.out <options> -pghpf -np #cpu

Where

**option**(s) One or more command-line options. For a *list* of options, use the **pghpf** option **-help**. For a listing and *description* of the **pghpf** specific compiler command-line options, refer to *pghpf User's Guide*, Chapter Two, "Compiler Options."

**filename** is the name of the source file, assembly-language file, object file, or library to be processed by the compilation system. More than one filename may be specified.

## pghpf File Extension Summary

| **Filename Extension** | **Assumed Type** |
|---|---|
| `.f`, `.hpf`, `.for`, `.f90` | HPF source programs (**pghpf** accepts the `.f90` extension) |
| `.F` | HPF source programs to be preprocessed |
| `.s` | Symbolic assembly language source |
| `.o` | Object files |
| `.a` | Libraries, linked if requested |
| other | Object files, passed to linker if requested, warning |

# 4. Message Passing Models

## Overview

There are two portable, abstract models of parallel execution that are supported on the O2K. Each provides a method of distributing a computation within a single-memory system or across the nodes of a multiple-memory system, without having to reflect the system configuration in the source code. These two programming models are Message Passing Interface (MPI) and Parallel Virtual Machine (PVM).

## Message Passing Interface Model

MPI is a portable standard programming interface for the construction of portable, parallel applications in C, FORTRAN 77, or FORTRAN 90. In a MPI program, insure that the following actions are taken:

- The source code includes the command ***INCLUDE*** **"mpif.h"** if written in FORTRAN 77 or FORTRAN 90, or **#include "mpi.h"** if written in C.

- On the compile line, use the **-lmpi** option.

- To run a MPI program, use the command

  ```
  mpirun -np N a.out [user_arguments]
  ```

  or

  in csh or tsh

  ```
  setenv MPI_NP N
  a.out [user_arguments]
  ```

  in ksh

  ```
  export MPI_NO N
  ```

  where ***N*** is the number of processes to start.

For on-line information about MPI, use **man mpi**, **man mpirun**, or execute the Insight Viewer (**insight**) and perform a search on the topic MPI.  This will return the *MPI User's Guide*,

*MPI Reference Manual*, and other documents that are available on the topic.  Additional detailed information about MPI can be found on the WWW at the URLs listed below:

**www.wag.caltech.edu/home-pages/tahir/local_mpi.html**
**www.erc.msstate.edu/mpi/index.html**
**www-jics.cs.utk.edu/PCUE/WWW/MOD7_MPI/index.htm**

## Parallel Virtual Machine Model

PVM is an integrated set of software tools and libraries that emulates a general-purpose, flexible, heterogeneous, concurrent computing framework on interconnected computers of varied architecture.  Using PVM, the programmer can create a parallel application that executes as a collection of concurrent processes on a set of computers. The set can include single O2K processors, multiple O2K processors, and nodes of Array systems.

To use PVM, the following actions should be taken:

- Set the environment variables listed below as shown:

  PVM_ROOT=/usr/array/PVM
  PVM_ARCH=SGI32mips4
  PVMDPATH=/usr/array/PVM/lib/pvmd

- The PATH environment variable needs to include the following directories:

  /usr/array/PVM/bin/SGI32mips4
  /usr/array/PVM/lib

- The source code includes the command *#include "/usr/array/PVM/include/pvm3.h"* (if written in C).

- The source code includes the command *include "/usr/array/PVM/include/fpvm3.h"* (if written in Fortran).

- On the compile line, use the following options:

  ***-lpvm3***

- One method of running PVM is to create a hostfile that contains the statements given below:

** wd=/u/user_name ep=$PVM_ROOT/bin/$PVM_ARCH:/u/user_name sdf1*

and then execute the job with the following commands:

```
pvmd hostfile &
a.out
```

For on-line information on PVM, use **man pvm**, **man pvm_intro**, **man pvmd3**, or execute the Insight Viewer (**insight**) and perform a search on the topic PVM. This will return the PVM documents that are available. Additional detailed information about PVM can be found on the WWW at **www.epm.ornl.gov/pvm**.

# 5. Programming Environment

## Batch Request Environment

The CEWES MSRC uses the Portable Batch System (PBS) for production job management on the Origin 2000 cluster. Jobs are queued from the 16-CPU interactive operating system image (named `origin`) and executed on the 122-CPU image (named `sdf1`). Jobs are scheduled for execution according to resources requested and current load levels on `sdf1`. Commands to submit jobs are similar to NQS commands, but see the **qsub** man page for details. Currently, the origin cluster consists of four queues.

1. Debug queue. The default walltime is 30 minutes. The maximum walltime is 30 minutes. There is no access control, so anyone with an account on the Origin cluster may submit a job to this queue. There is no restriction on the number of CPUs that can be requested in this queue.

2. Primary queue. The default walltime is 5 hours. The maximum walltime is 12 hours. There is no access control, so anyone with an account on the Origin cluster may submit a job to this queue. There is currently no restriction on the number of CPUs that can be requested in this queue.

3. Special queue. This queue is reserved for jobs requiring walltime greater than 12 hours. Users must request access to this queue and provide justification for such access. There is currently no restriction on the number of CPUs that can be requested in this queue.

4. Dedicated queue. This queue is reserved for high priority jobs that require the entire machine, or a large number of CPUs dedicated to a job. Its use is and should be extremely limited, for obvious reasons. Access to this queue is controlled, and access is gained by special request and sufficient justification.

To submit a job for execution on the O2K, first create a PBS script. Such a script might look like the following:

```
#!/bin/ksh
# This is a sample PBS script
#
```

```
# Export login shell environment to PBS
#PBS -V
#
# Merge stdout and stderr streams into a single output file
#PBS -j oe
# Ask for 48 CPUs
#PBS -l ncpus=48
#
# Request 1 hour of wallclock time for execution
#PBS -l walltime=01:00:00
#
# Change directory to home directory and copy
# input file to working storage
cd $HOMEDIR
cp input.dat $WORKDIR
#
# Change directory to working storage space
cd $WORKDIR
#
# Check that mass storage facility is available;
# if not, wait in quiet mode (no output messages)
msfstat -q -w
#
# Copy file "test.exe" from MSF in user's subdirectory
"source"
msfget $MSFHOME/source/test.exe $WORKDIR
#
# execute program
mpirun -mp48 test.exe > test.out
#
# recheck availability of MSF before attempting to move data
msfstat -q -w
#
# copy output file to MSF
msfput test.out $MSFHOME/output/test.out
#
# clean up files no longer needed
/bin/rm test.exe test.out
```

To submit a script for execution, type

**qsub <script_name>**

PBS will respond with the ID number of the batch job.

To inquire after the status of the job, use the **qstat** command. To see a verbose status of the job, type

```
qstat -f <job_id>
```

Man pages are available for all PBS commands.

# Interactive Environment

Users may interactively issue a sequence of commands to process, post-process, and archive their data and computational results to the MSF. Users should not execute their program or job interactively on `origin.wes.hpc.mil`. Host origin is for compiling, editing, and general interactive use by all CEWES MSRC Origin cluster users. In the interest of all users logged in, jobs found running on host `origin` may be unilaterally terminated.

When interactive use becomes more limited, requests to increase disk space quotas or interactive limits may be submitted by contacting Customer Support for a special request form. The following information must be supplied by the user for evaluation of the request by the system administrators and CEWES MSRC management:

- the amount of system resource requested
- the length of time requested for the increase
- special deadlines for the project
- an explanation of the attempts to work within the limits